



CECIDS

Chicago Early Childhood
Integrated Data System

Building an Early Childhood Integrated Data System with Integrity

A Field Guide From Chicago

September 2026

Table of Contents

1. Why Chicago Needed an Integrated Data System

- 1.1 Chicago's Mixed-Delivery Early Childhood System
- 1.2 Fragmented Data Also Meant Distributed Authority
- 1.3 From Program Records to a System-Level Picture
- 1.4 What This Meant for CECIDS

2. Governance and Investment Created the Conditions for Shared Use

- 2.1 From Distributed Authority to Shared Governance
- 2.2 Data Governors, Stakeholders, and Decision Rights
- 2.3 Making Governance Operational
- 2.4 Investing in the Infrastructure Around the Platform

3. Building and Refining the Technical System

- 3.1 Turning Shared Ambition Into Technical Infrastructure
- 3.2 What Sustained Use Made Visible
- 3.3 Strengthening the Technical Model
- 3.4 Building an Architecture That Could Learn

4. Establishing Data Meaning as Technical Infrastructure

- 4.1 From Source Data to Shared Meaning
- 4.2 Learning, Validating, Confirming, and Maintaining
- 4.3 Documentation as Institutional Memory
- 4.4 Preserve Difference Before Standardizing It

5. Designing for Current and Historical Use

- 5.1 One System, Different Time Horizons
- 5.2 Supporting Current and Operational Use
- 5.3 Preserving a Historical and Analytic Record
- 5.4 Designing for Both the Present and the Record

6. From Use Cases to Evidence

- 6.1 From Possibilities to Priorities
- 6.2 Looking Beneath the Use Case
- 6.3 Start With the Decision and the Evidence Needed
- 6.4 Knowing What Administrative Data Can and Cannot Answer

7. Turning Integrated Data Into Useful Information

- 7.1 Products Begin With Utility, Not Visualization
- 7.2 From Information Need to Product

7.3 Different Audiences Require Different Product Rules

7.4 Products in Service of System Learning

8. Research as a Distinct Form of Data Use

8.1 When Administrative Analysis Becomes Research

8.2 Research Requires a Distinct Governance Pathway

8.3 Stewardship and Methodological Independence

8.4 Returning Research to the System

9. Trust as an Operating Property

9.1 Trust Was Built Through Operation

9.2 Making Reliability Visible

9.3 Different Stakeholders Need Different Assurances

10. Sustainability and Responsible Closeout

10.1 Sustainability Beyond Platform Permanence

10.2 Preserving Knowledge and Stewardship

10.3 Closeout as a Designed Process

10.4 Sustainability as Capacity That Can Travel

11. Conclusion: What Endures

11.1 CECIDS as Information Infrastructure

11.2 What the CECIDS Experience Offers the Field

11.3 What Survives the Transition

12. Companion Resource Set

12.1 Technical and Practice Memos

12.2 Templates and Operating Artifacts

12.3 Collaborative Works

12.4 Future and Closeout Resources

1. Why Chicago Needed an Integrated System

1.1 Chicago's Mixed-Delivery Early Childhood System

Chicago's early childhood system is intentionally designed as a mixed-delivery system, which brings strength and flexibility. It also produces structural complexity that makes system-level data integration inherently fragile. Across the city, public agencies, schools, Head Start recipients, community-based organizations, and other partners administer programs that differ in funding sources, eligibility rules, service models, reporting requirements, and accountability structures. Together, these programs create a broad network of early care and education options for children and families while also creating an information environment in which no single administrative system can describe the whole.

The diversity of the system is not incidental. Families may receive services through school-based preschool, community-based centers, home visiting, child care assistance, Head Start or Early Head Start, state-funded preschool and zero to three programs, early intervention, special education, or combinations of these supports. Community-based organizations may operate multiple sites and classrooms while braiding or blending several public funding streams. Public agencies and intermediaries may play different roles as funders, administrators, grantees, reporters, or regulators. This structure allows programs to respond to different family needs and community contexts, but it also means the system's data are generated through multiple institutional pathways rather than a single common operating model.

Each of those pathways produces legitimate information for a specific purpose. Programs collect and maintain data to determine eligibility, document enrollment and attendance, support billing and reimbursement, monitor services, meet grant or contract requirements, and comply with local, state, and federal reporting obligations. Those data are highly useful within the program or funding stream that created them. The challenge arises when the questions extend beyond a single program: How many unique children are being served across Chicago? Where are they being served? Through which combinations of programs and funding? Where do participation patterns differ across communities? What does the available supply of services look like relative to the population of young children?

Key Principle: A mixed-delivery system can be both programmatically strong and analytically fragmented. Integration should preserve the strengths and legitimate differences of the system while creating enough shared structure to support cross-system learning.

1.2 Fragmented Data Also Meant Distributed Authority

The fragmentation of Chicago's early childhood data reflected a broader institutional reality: authority was also distributed. No single organization controlled the early childhood system from end to end, and no single entity held all of the legal, operational, and contextual knowledge required to define systemwide meaning on its own.

The organizations contributing data to the Chicago Early Childhood Integrated Data System (CECIDS) carried responsibilities to children and families, funders, regulators, boards, and communities. They

understood how their programs operated and would often bear the consequences if an integrated analysis mischaracterized those programs. At the same time, policymakers, funders, researchers, advocates, and families had legitimate needs for information that crossed organizational boundaries. A citywide system therefore had to create a way for those interests to coexist without allowing technical control of the platform to become de facto authority over data meaning and use.

That distinction is central to the CECIDS story. Technical integration can connect records, but it cannot by itself determine who has the authority to approve a use, define a shared measure, resolve conflicting interpretations, or stand behind a public conclusion. Those questions require governance. For CECIDS, the need for integration and shared governance was therefore inseparable from the beginning.

1.3 From Program Records to a System-Level Picture

CECIDS therefore faced a different challenge than simply assembling more data: it had to determine when records from different programs represented the same child, service relationship, organization, or point in time—and when they did not. Administrative records that were accurate within their original program context could not automatically be combined into a system-level count without first understanding the relationships among them.

A child may legitimately appear in more than one administrative dataset during the same year because each program is responsible for documenting its own service relationship. The same provider may report different information to different agencies, and the organization receiving a grant may not be the same organization operating the classroom or reporting the child-level record. In this environment, duplicate records are not simply errors to be removed; they may represent distinct and valid program relationships that must be understood before they are reconciled.

The same challenge applies to organizational and geographic relationships. Providers can operate multiple sites, sites can host multiple program types, and program administration may sit at a different organizational level from service delivery. If those relationships are flattened too early, an integrated system can produce counts that are technically consistent but operationally misleading. The task is therefore not merely to deduplicate records or standardize names. It is to reconstruct enough of the underlying relationships to understand what each record represents and when records should, and should not, be combined.

This is why the phrase “data-rich but insight-poor” became useful in describing the pre-integration environment. The issue was not a lack of information. It was the difficulty of assembling that information into a defensible system-level picture without losing the context that made the source data meaningful.

1.4 What This Meant for CECIDS

CECIDS emerged as an effort to build shared information capacity across a system that was intentionally distributed. Its purpose was not to replace program-specific systems or eliminate the distinctions among them. Instead, it was to create an infrastructure that could connect information across those systems,

support clearer system-level questions, and do so in a way that respected the responsibilities of the organizations contributing the data.

That required more than a common repository. It required governance, legal agreements, technical architecture, provider engagement, business analysis, documentation, product development, and research processes that could work together. Those functions did not mature all at once, and the system refined them as new source systems, use cases, and questions revealed additional complexity. The value of the CECIDS experience lies in that accumulated learning: the project increasingly understood integration not as the elimination of difference but as the disciplined management of difference in service of shared understanding.

Key Principle: The central challenge of an integrated early childhood data system is not making unlike datasets look the same. It is creating a shared representation that is useful across the system while preserving the relationships, definitions, time periods, provenance, and limitations that give the source data meaning.

2. Governance and Investment Created the Conditions for Shared Use

2.1 From Distributed Authority to Shared Governance

CECIDS was built within a system in which authority was distributed among organizations, each with distinct legal obligations, operational responsibilities, and institutional accountabilities. In that environment, governance was not an administrative layer added after technical development. It was the mechanism through which participating organizations established how decisions would be made about data use, access, standards, products, research, and public reporting.

The practical need for governance was straightforward. Once information from multiple organizations was brought together, the system needed an explicit way to determine which uses were appropriate, how shared definitions would be established, who could access sensitive information, what could be released publicly, how new analyses would be reviewed, and how disagreements would be resolved. Without an explicit governance structure, those decisions would still occur, but they would tend to migrate to informal channels or default to whichever organization held the technical infrastructure, funding, political authority, or staffing capacity at a particular moment.

CECIDS chose instead to make those responsibilities visible. The CECIDS Governance Agreement established formal roles, decision rights, and procedures that reflected the distributed nature of the early childhood system. This structure created a common framework for decision-making while preserving the stewardship responsibilities of participating organizations.

2.2 Data Governors, Stakeholders, and Decision Rights

The CECIDS Governance Agreement established a two-tier Executive Committee that distinguished between organizations with formal stewardship responsibility for contributed data and a broader set of

stakeholders with a legitimate interest in how the system developed and how integrated information would be used. The Data Governors Tier included participating organizations responsible for contributing and governing data. The Stakeholders Tier broadened the table to include public agencies, provider representatives, philanthropy, researchers, family and community voices, and other early childhood stakeholders.

This distinction was structural rather than symbolic. The Data Governors held formal decision rights over specified matters, including the selection and management of a Technical Administrator, standards and business rules, data access and use procedures, final data products and use cases, and procedures for review of research, reports, and analyses using CECIDS data. The broader Executive Committee contributed to priorities, stakeholder engagement, research agendas, and the incorporation of the system's guiding principles, but it did not hold the same binding authority as the organizations contributing to and standing behind the data.

This operating model reflected a central lesson about distributed systems: meaningful participation does not require identical authority. Different stakeholders bring different forms of expertise, responsibility, accountability, and risk. A provider may understand how enrollment is operationally recorded. A public agency may carry regulatory or administrative authority. A researcher may bring methodological expertise. Families and advocates may identify interpretive or equity concerns that would not otherwise be visible. A legitimate governance structure makes these roles explicit rather than treating them as interchangeable.

The governance model also preserved specific individual Data Governor rights. Each Data Governor remained responsible for determining the sensitive information it submitted and the conditions under which that information could be accessed within CECIDS, subject to governing agreements and applicable law. Data Governors also retained rights to review and comment on research, reports, or analyses using the data they had contributed before publication under established procedures.

Key Principle: Meaningful participation does not require identical authority. A legitimate governance structure makes different forms of expertise, stewardship, accountability, and decision-making responsibility explicit.

2.3 Making Governance Operational

CECIDS also separated governance authority from day-to-day technical administration. Northern Illinois University (NIU) served as the Technical Administrator and operated the technical environment through separate individual agreements with participating Data Governors. The shared governance structure determined approved uses, standards, access conditions, and products, while the Technical Administrator was responsible for implementing those decisions through the technical system and associated operating processes.

This separation supported the distributed character of the system. It avoided making the City of Chicago the default owner of all integrated information, reduced the risk that one provider or agency would control the system simply because it operated the platform, and created clearer legal relationships between

participating organizations and the entity responsible for technical administration. The model also reinforced the principle that technical control of the environment should not automatically become authority over data meaning or use.

The arrangement required sustained coordination because governance decisions had to be translated into technical implementation. Standards, access permissions, approved use cases, publication rules, and research conditions were not meaningful if they remained only on paper. Project leadership and technical staff therefore had to maintain a continuous connection between governance decisions and the way the system actually operated. Governance established legitimate authority, but technical administration made that authority operational.

Governance became meaningful through ordinary decisions. A proposed product could be narrowed because the data did not support the intended interpretation. A new use could require additional review. A definition could be revisited because contributing organizations understood the same concept differently. In some cases, definitions were co-designed with data-contributing agencies because the construct itself was highly specific to the program or operational context and could not be responsibly inferred from the data structure alone. In each case, the system needed a way to identify who held authority, who needed to be consulted, who was responsible for implementation, and how a concern could be raised.

Over time, clear decision rights reduced uncertainty. Once roles, review pathways, and escalation processes became familiar, participants did not have to renegotiate authority every time a new question arose. Predictability became one of governance's most important benefits. The value of mature governance was not that it removed disagreement; it made the path for resolving disagreement more predictable.

2.4 Investing in the Infrastructure Around the Platform

CECIDS' development also demonstrated that the technical platform was only one part of what had to be funded. The system required sustained investment in planning, legal agreements, governance support, provider engagement, project leadership, business analysis, documentation, architecture, engineering, analytics, research processes, and stakeholder communication.

Much of that work was largely invisible to end users. A governance meeting does not look like a dashboard. A data-sharing agreement does not produce a new metric. A provider conversation about how a data field is used locally may not create an immediate public product. Yet these activities determine whether later products can be interpreted, defended, and maintained.

Public leadership and philanthropy played an important role because they created space for work whose value was cumulative rather than immediately visible. Patient investment allowed CECIDS to develop governance and technical structures that could support participation, learning, and refinement over time. It also allowed the system to absorb the time required for legal review, relationship-building, technical documentation, and iterative validation rather than treating those activities as delays to the "real" work.

The experience further demonstrated that governance roles themselves have to be resourced. Decision rights on paper are not enough. Someone has to prepare materials, track decisions, translate technical issues for

decision-makers, document resolutions, follow up with organizations, maintain agreements, and ensure that approved processes are reflected in technical operations. These functions are not administrative extras; they are part of the system's interpretive and institutional infrastructure.

Key Principle: In a distributed early childhood system, governance creates the authority, predictability, and accountability required for shared data use. Those functions are infrastructure and must be designed, operated, and funded as such.

3. Building and Refining the Technical System

3.1 Turning Shared Ambition Into Technical Infrastructure

The initial CECIDS build translated years of planning, governance development, legal work, and stakeholder engagement into a functioning technical environment. It created the first infrastructure for receiving and integrating data from multiple early childhood organizations and demonstrated that a shared citywide system was technically and institutionally possible. That foundation moved CECIDS from aspiration to implementation and, importantly, created the conditions for the system to learn from real data, real use cases, and real operating variation.

The early architecture also reflected the information and source-system environment available at the time. Several participating organizations used similar systems, and the initial design reasonably responded to the proof-of-concept needs of the project. As CECIDS began operating across more sources, more years, and more use cases, the system gained a deeper understanding of the variation that would need to be represented and maintained over time.

3.2 What Sustained Use Made Visible

As CECIDS integrated data from more organizations, accumulated additional years of information, and supported more demanding questions, the system developed a deeper understanding of the variation embedded in source systems and program operations. Provider systems changed, local configurations differed, and the same commercial platform could be used in materially different ways across organizations. Use cases also became more specific, increasing the need to understand not only whether a field could be ingested, but what operational relationship the field actually represented.

As use expanded, CECIDS encountered differences that were not visible from schemas or transmission formats alone. Fields that appeared similar could reflect different program relationships, local configurations, or reporting practices, while changes in source systems or policy could alter how the same concept was represented over time. The technical challenge therefore became increasingly semantic as well as structural.

Sustained use also increased the importance of historical interpretation. A value that was correct in one program year might be represented differently after a source-system change or policy revision. As additional

years accumulated, the system needed stronger conventions for effective dates, versioning, source provenance, and documentation so that historical data could remain interpretable over time.

3.3 Strengthening the Technical Model

By 2024, CECIDS had accumulated enough operational experience to undertake a significant technical restructuring. The work provided an opportunity to strengthen the alignment among source data, provider operations, integrated data models, cleansing and transformation logic, funding relationships, participation definitions, and documentation.

The purpose of the restructuring was not simply to replace technical components. It was to make the relationships among program operations, source-system representation, and integrated-system logic more explicit and maintainable. In practice, this meant revisiting assumptions, documenting business rules in greater detail, improving traceability between source interpretations and technical implementations, and strengthening the architecture's ability to accommodate variation across providers and over time. That stronger alignment increased the system's ability to distinguish genuine cross-system patterns from differences created by source configuration, program rules, or reporting practice.

The restructuring built on the work that preceded it. Years of technical implementation, governance, provider participation, and use-case development gave CECIDS a richer understanding of the operating environment than was available during the initial build. That accumulated knowledge made it possible to strengthen data models, transformation logic, documentation, and traceability while designing more explicitly for variation across providers and over time.

3.4 Building an Architecture That Could Learn

As CECIDS' understanding of its source systems deepened, the project also made the connection between operational knowledge and technical design more explicit. Business analysis, provider engagement, documentation, and technical implementation increasingly operated as connected functions rather than separate activities. New information about a source system could therefore lead to review of mappings, transformations, quality rules, measures, or historical interpretation.

This mattered because the technical environment was not static. Source systems changed, program rules changed, staff changed, and new analytic questions exposed assumptions that earlier uses had not tested. A maintainable architecture therefore needed not only technical flexibility, but mechanisms for preserving institutional memory and incorporating new knowledge without losing traceability to earlier decisions.

The next section examines how CECIDS made that learning more systematic by connecting provider knowledge, business analysis, validation, documentation, and technical implementation.

Key Principle: Architecture evolves as understanding evolves. A mature integrated data system preserves enough technical flexibility, provenance, documentation, and institutional memory to incorporate new knowledge without losing the meaning of what came before.

4. Establishing Data Meaning as Technical Infrastructure

4.1 From Source Data to Shared Meaning

By the time CECIDS had strengthened its technical model, one lesson had become increasingly clear: an architecture that can adapt also needs a disciplined way to learn what its source data mean. Files, schemas, application programming interfaces, and vendor documentation can describe how information is structured and transmitted. They do not always explain the operational event that caused a value to be entered, the business rule that shaped it, the staff practice that maintained it, or the local configuration that gave it meaning.

That distinction mattered in CECIDS because participating organizations often operated under different program rules, funding requirements, reporting expectations, and workflows, even when they used similar systems. A field could be transferred correctly and still be interpreted incorrectly. If that interpretation then became embedded in a transformation, measure, analytic model, or product, a technically valid pipeline could produce a substantively misleading conclusion.

CECIDS therefore treated source meaning as part of the technical problem itself. Understanding how programs operated, how source systems were configured, and how staff used particular fields became necessary for determining what could be integrated, what needed to remain distinct, and what could responsibly support a shared measure. The goal was not to privilege informal explanation over technical evidence. It was to create a more complete evidentiary base for technical design.

4.2 Learning, Validating, Confirming, and Maintaining

Over time, CECIDS developed a more structured operating cycle for building that understanding:

Learn → Validate → Confirm → Maintain.

The sequence reflected an important distinction between hearing an explanation and establishing a durable interpretation. Operational knowledge was essential, but it remained working evidence until it had been tested against other available sources, including program policy, source documentation, observed data, technical artifacts, and implementation logic.

Structured Data Dialogues and business analysis became mechanisms for carrying out this work. They created a repeatable way to ask how data were created and used, investigate apparent discrepancies, identify exceptions and local practices, and return emerging interpretations to the people responsible for the source for confirmation. In some cases, the underlying construct was specific enough to the program or operational context that its meaning could not be responsibly inferred from the data structure alone.

The technical burden, however, remained with CECIDS. Program and provider staff were not expected to translate their operations into code or redesign integration logic. CECIDS was responsible for asking questions in the language of program operations, reconciling the answers with technical evidence,

translating confirmed understanding into mappings or transformations, and returning the resulting interpretation in a form the source organization could review.

This process also changed how discrepancies were treated. An unexpected count or unfamiliar value could indicate a source error, but it could also reflect a legitimate difference in definition, reporting period, system configuration, or workflow. Investigating meaning before classifying a discrepancy as a data-quality problem allowed CECIDS to distinguish technical defects from operational variation and to address each appropriately.

4.3 Documentation as Institutional Memory

As CECIDS accumulated more years of data and more detailed source knowledge, documentation became the mechanism for keeping that understanding from remaining dependent on individual memory. The challenge was not simply to create more documents. It was to decide which kinds of knowledge needed to be preserved, how durable each finding was, and where that information should live so that it could be maintained and used later.

CECIDS developed a connected documentation set for this purpose. Working notes captured questions, evidence, and unresolved issues. Confirmed structural information was maintained in Provider Fact Sheets. Program-year changes, exceptions, and interpretation issues were preserved in Annual Data Summaries, while Yearly Overviews recorded the programs, sites, relationships, and operating footprint active during a particular year. Technical implementation decisions remained in ETL and mapping documentation. These artifacts served different functions, but together they helped preserve the relationship between program operations, technical representation, and analytic interpretation.

The distinction among structural, year-specific, and technical knowledge was especially important. A temporary practice should not become permanent architecture simply because it appeared in one year of data, and a technical workaround should not be mistaken for a program rule. Preserving effective periods and prior versions also allowed CECIDS to maintain historical interpretability rather than rewriting earlier records to match current conditions.

Documentation therefore functioned as institutional memory. It reduced dependence on individual staff, supported onboarding and technical continuity, and made it easier to understand why a rule, mapping, or interpretation existed. Just as importantly, it created a place for uncertainty to remain visible until it had been resolved, rather than forcing premature standardization.

4.4 Preserve Difference Before Standardizing It

A mature integrated system eventually has to create common representations, but CECIDS learned that standardization should follow understanding rather than substitute for it. Similar-looking data elements can represent different concepts across programs, source systems, units of analysis, and time periods. The task is not to eliminate those distinctions prematurely, but to determine which differences matter for the intended use.

Funding, capacity, and race and ethnicity data each illustrated this broader principle. Funding information can describe awards, allocations, classroom designations, child associations, or payments. Capacity can refer to licensed, physical, funded, staffing, or operational conditions. Race and ethnicity data may reflect different collection standards, category structures, and levels of detail across time. In each case, collapsing differences too early can create false equivalence or false precision.

CECIDS therefore increasingly distinguished among source values, integrated representations, and derived measures. Source information needed to remain traceable to its origin. Integrated representations needed to make cross-source use possible without erasing meaningful distinctions. Derived measures required explicit definitions, assumptions, and limitations appropriate to the decision or analysis they were intended to support.

That discipline connected technical design directly to decision quality. A common measure is useful only if users can understand what it represents and what it does not. Integration adds value when it makes relationships visible across sources without creating a level of certainty that the underlying data cannot support.

Key Principle: A schema describes structure; it does not establish business meaning. Integrated data become trustworthy when source meaning is learned, tested, documented, and maintained—and when meaningful differences are preserved rather than flattened.

5. Designing for Current and Historical Use

5.1 One System, Different Time Horizons

Once CECIDS could represent source meaning more reliably, the next question was what the information infrastructure needed to preserve for different kinds of use. Not every decision operates on the same time horizon. Some users need current or regularly refreshed information about what is happening now. Others need a durable historical record that supports trend analysis, population estimates, policy questions, or research.

CECIDS had to support both. Provider-facing and authorized citywide products depended on relatively current information that could be refreshed, reconciled, corrected, and interpreted in an operational context. Historical and analytic work depended on preserving prior records, definitions, effective periods, transformation logic, and source provenance so that the system could explain not only what was known, but how that knowledge had been constructed over time.

These were not competing purposes. They were different expressions of the same broader goal: improving the information available for decision-making and system learning. The architectural challenge was to support one without weakening the other.

5.2 Supporting Current and Operational Use

Current and operational uses placed a premium on predictability. Data had to arrive on a cadence that matched the intended use, be ingested and reconciled consistently, and be available to the appropriate audience under the right access and privacy conditions. If a discrepancy was identified, users needed a clear pathway for correction and follow-through. If a source changed, the product logic and documentation had to change with it.

CECIDS supported multiple product contexts, including provider-facing information, authorized citywide views, and public-facing products. These uses drew from related underlying data but required different rules for access, aggregation, reconciliation, interpretation, and disclosure. A provider may need detailed information to reconcile its own records. An authorized system leader may need cross-provider information to understand citywide patterns. A public product requires a level of aggregation and explanation appropriate for broad release.

The cadence of the source data also mattered. CECIDS did not treat “real time” as a universal standard because many administrative sources were not created or transmitted in real time. What mattered was whether the refresh schedule was sufficiently current for the decision the product was intended to support, and whether users understood the period represented.

Operational usefulness therefore depended on more than speed. It depended on a reliable relationship among source cadence, ingestion, reconciliation, access, interpretation, correction, and maintenance. A current product that cannot be explained or corrected may be less useful than a slightly older product whose meaning and limitations are clear.

5.3 Preserving a Historical and Analytic Record

Historical and analytic uses required a different kind of discipline. CECIDS needed to preserve information in ways that allowed analysts and researchers to reconstruct what was known during a given period and to compare years without silently applying current definitions to older data. This required attention to effective dates, historical retention, versioned definitions, reproducible transformations, record linkage, provenance, and program-year conventions.

The value of that preserved record extended beyond traditional research. It supported unique-child estimates, participation and overlap analyses, historical trends, geographic analysis, population estimates, and policy questions that could not be answered from a current snapshot alone. These uses depended on being able to distinguish a genuine change in the early childhood system from a change in how a source collected, represented, or transmitted information.

CECIDS’ documentation practices reinforced this work. Structural information was separated from year-specific conditions, and historical records were not “cleaned up” simply to make earlier years conform to current practice. Preserving what was true, known, and documented at the time protected the interpretability of longitudinal analysis.

The same principle applied to research-ready datasets. A durable analytic asset requires more than linked records. It requires enough documentation and provenance for future users to understand the population represented, the rules used to construct measures, the periods covered, and the limitations that should shape interpretation.

5.4 Designing for Both the Present and the Record

The tension between current and historical use is easy to underestimate because both may rely on the same underlying sources. An architecture optimized only for current dashboards may overwrite prior values, simplify effective dates, or prioritize the latest state of a record. An architecture optimized only for retrospective analysis may preserve history well but fail to provide information on a cadence that supports operational decisions.

CECIDS' experience reinforced the need to design deliberately for both. Current use depends on timely and maintainable information. Historical use depends on preserving context, versions, and provenance. Both depend on clear definitions and stewardship, but they place different demands on how records are stored, transformed, documented, and accessed.

The reason to preserve both time horizons is not technical elegance. It is to expand the range of legitimate questions the system can support. A city needs to know what is happening now, but it also needs to understand how patterns changed, whether apparent differences are comparable across time, and what earlier conditions can teach about future decisions.

Key Principle: An ECIDS must be designed for both the present tense and the historical record. The value of integration depends not only on what the system can show now, but on whether it can explain how information was represented, defined, and used over time.

6. From Use Cases to Evidence

6.1 From Possibilities to Priorities

CECIDS' early use-case development reflected the breadth of what stakeholders hoped integrated data could make possible. During planning, stakeholders articulated 52 distinct use cases spanning participation, population, resource allocation, system learning, and other questions that could not be answered easily within a single program or funding stream. The catalogue gave CECIDS a concrete picture of stakeholder demand for information and helped connect system design to questions that mattered across the early childhood community.

As the system matured, the challenge shifted from identifying possibilities to deciding which questions could be answered well enough to support action. Prioritization became increasingly important because each use case carried requirements for definitions, source coverage, business rules, analytic methods, governance, and product design. A broad list of potential questions could guide ambition, but it could not

by itself establish which questions were ready for analysis or what level of confidence the underlying data could support.

CECIDS therefore began working more deliberately from prioritized questions back toward the data and definitions required to answer them. This was not a rejection of the original use cases. The initial catalogue created the inventory from which later discipline could develop. Sustained use made it possible to distinguish questions that were technically feasible from questions that were sufficiently defined, actionable, and interpretable to justify continued investment.

6.2 Looking Beneath the Use Case

A deeper challenge emerged even after use cases had been prioritized: the requested metric or dashboard was not always the same as the underlying information need. By 2025, CECIDS partnered with NORC at the University of Chicago to examine whether prioritized population and participation use cases were adequately capturing what decision-makers actually needed to know.

That work showed that use cases can function as proxies for more complex questions. A request for an enrollment count may reflect concern about workforce constraints, eligibility, capacity, funding, or family access. A participation gap may matter because of what it implies about where services are available, who can use them, or what barriers shape enrollment. Clarifying the underlying need therefore required moving beyond the requested metric to ask what decision or action the information was meant to inform.

The NORC work helped CECIDS articulate a set of broader North Star Questions and distinguish between descriptive questions that administrative data could plausibly address and more structural questions that would require new information, research design, or additional forms of evidence. The lesson was not that use cases were insufficient. It was that they were one tool within a larger process for understanding policy and program information needs.

This shift also reinforced the importance of continued iteration. Information needs change as programs change, new evidence becomes available, and decision contexts evolve. A mature integrated system therefore cannot assume that the questions defined during planning will remain the only or best questions throughout the life of the system.

6.3 Start With the Decision and the Evidence Needed

CECIDS increasingly worked backward from the decision rather than forward from the available data. A useful starting point was to clarify the information need and the action it might inform, then specify the population, geography, time period, subgroup, and definition required to make the question meaningful. Only then could the system assess what evidence would be necessary and whether the available administrative data could provide it.

This process also forced greater attention to actionability. One practical question was whether materially different answers would change what a stakeholder planned to do. If a participation rate of 14 percent

would lead to the same decision as a rate of 60 percent, the information need might not yet be defined clearly enough to justify additional analytic investment. If the response would differ substantially, the required level of precision and timeliness could be considered in relation to the decision at hand.

The same reasoning applied to the cost of producing more precise or more current information. Greater precision may require additional data collection, more complex linkage, more frequent source feeds, or more intensive validation. Those investments are valuable when they materially improve a decision, but not every question requires the maximum achievable level of detail.

The result was a more policy-facing sequence:

information need → intended decision or action → necessary definitions → actionability → evidence required → feasibility and limitations.

This sequence kept the system from treating every available field as an invitation to produce a metric.

6.4 Knowing What Administrative Data Can and Cannot Answer

One of the most important capabilities CECIDS developed was the ability to distinguish among different kinds of answerability. Some questions could be answered directly from integrated administrative data. Others could be answered only with explicit qualifications. Some required new data collection, qualitative information, surveys, or external sources. Still others were better suited to research because they involved causal, developmental, or structural questions that descriptive administrative data could not resolve on their own.

Recognizing those boundaries was part of analytic maturity rather than a limitation of the system. Integrated administrative data can create powerful visibility across programs and populations, but they do not automatically explain why a pattern exists or what intervention would change it. A defensible system needs to be able to say when the available evidence is descriptive, when it is incomplete, and when a different form of inquiry is required.

That distinction also protected against false confidence. Producing a number because it can be calculated does not mean the number answers the question a policymaker or provider is asking. CECIDS became more valuable when it could distinguish a metric from evidence and identify when the gap between them required additional work.

Key Principle: Decision-makers need evidence, not merely metrics. A mature integrated data system knows both what it can answer and what it should not claim to answer.

7. Turning Integrated Data Into Useful Information

7.1 Products Begin With Utility, Not Visualization

Once CECIDS had a clearer understanding of the questions stakeholders were trying to answer, product development could begin from purpose rather than format. A dashboard, report, dataset, or analytic brief is not itself the outcome of an integrated data system. Its value depends on whether it improves the information available for a decision, supports shared understanding, reveals a meaningful pattern, or enables responsible analysis.

CECIDS' Theory of Action provides the appropriate frame. Integrated data were intended to support better decisions, services, advocacy, equity, and system learning. Products therefore needed to be understood as mechanisms within that larger theory, not as evidence of success simply because they had been built or released.

A useful product question was therefore not “What can we visualize?” but “What decision, action, understanding, or conversation should become better because this information exists?” That distinction helped connect product development to the information-needs discipline described in the previous section and prevented technical feasibility from becoming the primary reason to build.

7.2 From Information Need to Product

CECIDS' mature product logic can be summarized as a sequence:

Theory of Action → Stakeholder → Information Need → Intended Decision or Use → Evidence Required → Feasibility and Limitations → Product Form.

The product form comes late in the sequence because a dashboard, table, dataset, or narrative brief should follow from the nature of the question and audience rather than precede them.

Development did not end once a format had been selected. Prototypes needed to be interpreted and validated with the people who understood the source data and intended use. Governance determined whether and how the product could be released. Usage and feedback then informed whether the product should be maintained, revised, narrowed, or retired.

This lifecycle also meant that product development could legitimately end with a decision not to build. A product may be technically feasible but still provide insufficient value, create unacceptable interpretive risk, require a level of maintenance disproportionate to its use, or depend on evidence the system cannot support responsibly. Deciding not to proceed can therefore be an appropriate product decision.

The discipline is similar to the one CECIDS developed for use cases: start with purpose, make the evidence and limitations explicit, and allow the form of the output to follow.

7.3 Different Audiences Require Different Product Rules

CECIDS served different audiences through different product contexts. Provider-facing information, authorized citywide products, and public-facing views could draw from related integrated data, but they could not be governed or interpreted as though they were the same product.

Provider-facing information could support detailed reconciliation and operational understanding because the provider already holds the underlying source relationship. Authorized citywide products could enable broader system analysis while still operating within defined access conditions. Public products required stronger aggregation, disclosure review, accessibility, and explanatory context because they were designed for an audience without the same source knowledge or governance relationship.

These distinctions meant that broader access was not inherently better access. Appropriate access depended on purpose, audience, privacy, stewardship, and the level of interpretation the product could support. A measure suitable for internal reconciliation or support program monitoring might be inappropriate for public comparison. A citywide pattern may be useful for planning even when the underlying information should not be exposed at the provider or child level.

Product rules therefore had to account for more than technical permissions. They also reflected the degree of reconciliation required, the interpretive context available to the audience, the risk of misreading a measure, and the system's obligation to protect sensitive information while still supporting legitimate use.

7.4 Products in Service of System Learning

CECIDS' product experience ultimately returned to the same question that shaped its Theory of Action: did the information help the early childhood system understand itself more clearly and act more responsibly? A product that answered a real information need could support planning, shared interpretation, policy discussion, provider reconciliation, public understanding, or research. A product that did not connect to a meaningful use could become a maintenance burden regardless of its technical quality.

That made feedback and maintenance part of product design rather than post-launch activities. Definitions could change, source systems could change, stakeholder needs could change, and products could lose relevance. Maintaining a product therefore required the same disciplines that supported the rest of CECIDS: source understanding, documentation, governance, communication, and clear ownership.

The broader implication is that integrated data create potential, but products create value only when they are connected to legitimate questions and sustained use. The number of dashboards or reports produced is therefore a weak measure of system maturity. A smaller set of products that are understood, governed, maintained, and used can create more durable value than a larger portfolio that lacks clear purpose.

Key Principle: A product can be technically accurate and still be unusable, misleading, or untrusted. Product development begins with purpose and ends with sustained utility.

8. Research as a Distinct Form of Data Use

8.1 When Administrative Analysis Becomes Research

Not every consequential question can be answered through descriptive administrative analysis alone. Integrated data can show who is being served, where gaps exist, and how patterns change over time. Questions about why those patterns exist, what factors shape them, or whether a policy or program change produced a particular outcome may require a distinct research design and additional forms of evidence.

The distinction is driven by the question and the claim being made, not simply by analytic complexity. Research was therefore treated as a distinct form of data use within CECIDS—one that required its own governance, access, and review pathway.

8.2 Research Requires a Distinct Governance Pathway

Research required a distinct governance pathway because approval to conduct a study and authorization to use CECIDS data were related but different responsibilities. Where applicable, researchers remained responsible for obtaining approval through their Institutional Review Board or other appropriate institutional research process. CECIDS had an additional responsibility as steward of the integrated data: determining whether and under what conditions a proposed study could use CECIDS data in ways consistent with the authority, principles, and responsibilities established by the Data Governors.

The Governance Agreement already assigned Data Governors authority over standards and business rules, procedures for data access and use, and procedures for review and comment on research prior to publication. Earlier CECIDS policies required individual Data Governors to review research requests. As research use became more concrete, CECIDS recognized that preserving this authority did not require every Data Governor to independently review every study.

The Research Review Board translated Data Governor authority into a standing research-governance function. The Data Governors authorized the RRB to act on their behalf and in the best interests of the collective when reviewing, approving, and overseeing research involving CECIDS data. The RRB could assess whether a proposal was feasible with CECIDS data and capacity, aligned with CECIDS priorities and governance, compliant with applicable requirements, adequately protected, and valuable to Chicago's early childhood community.

In this sense, the RRB operated alongside rather than in place of institutional research review. Institutional research processes address the protections and requirements associated with the study itself. CECIDS research review addressed a related but distinct question: should this research be allowed to use CECIDS data, under what conditions, and does it meet the principles and responsibilities established by the Data Governors?

8.3 Stewardship and Methodological Independence

Research made the distinction between stewardship and methodological authority especially visible. Data Governors and data-contributing organizations had legitimate responsibilities related to privacy, factual accuracy, source context, and the ways their programs or communities might be characterized. Researchers, in turn, needed sufficient independence to design methods, analyze evidence, and report findings without allowing institutional preference to determine the result.

CECIDS treated those responsibilities as complementary rather than interchangeable. Contextual review could help identify factual errors, missing operational context, limitations not visible in the dataset, or interpretations that exceeded what the source information could support. That review strengthened the quality of the work because administrative data often carry meanings that are not fully visible through the analytic file alone.

De-identification remained important, but it did not end governance responsibility. A dataset without direct identifiers can still produce findings with implications for specific populations, communities, providers, or public programs. Responsible research governance therefore had to address not only confidentiality but also intended use, public value, analytic approach, contextual expertise, dissemination, and the potential consequences of interpretation.

At the same time, contextual review was not intended to provide a mechanism for suppressing methodologically sound findings. Researchers remained responsible for methodological integrity, and data contributors retained a defined role in stewardship and interpretation rather than authority over the analytic conclusion. This reflected the same governance principle established earlier in CECIDS: meaningful participation does not require identical authority.

8.4 Returning Research to the System

The value of research to an integrated data system extends beyond publication. Findings can test assumptions, reveal information gaps, refine definitions, and generate new questions that shape the system's future evidence agenda. Research can therefore be both an output of information infrastructure and a source of new learning for it.

CECIDS also distinguished between broader field-building research and analysis intended to support implementation or near-term adjustment. Both can be rigorous, but they serve different purposes and may require different timelines and products. In either case, findings are most valuable when their conclusions, limitations, and implications return to policymakers, providers, advocates, and other stakeholders in forms they can understand and use.

Key Principle: Research governance should protect both legitimate stewardship and methodological independence. A mature integrated data system creates a clear pathway for determining when its data may be used for research, under what conditions, and how findings return to the system without transferring control of the research conclusion to data stewards.

9. Trust as an Operating Property

9.1 Trust Was Built Through Operation

CECIDS began with a strong commitment to stakeholder participation and shared governance, but the project also learned that those commitments did not automatically produce trust in day-to-day operation. Early engagement could establish shared ambition and legitimate roles, but it could not anticipate every question that would emerge as data were transferred, interpreted, combined, and turned into products. At times, differences in expectations, interpretation, timelines, or validation made visible where relationships and operating practices needed to become more deliberate.

Those moments were part of the system's maturation. CECIDS increasingly understood that trust depended less on assurances than on what happened after a file arrived, a discrepancy surfaced, a commitment changed, or a stakeholder raised a concern and more on whether the system acknowledged the issue, investigated it with the people who understood the source, communicated candidly about what was known and unknown, documented the resolution, and followed through. Over time, those repeated practices became as important to system credibility as the technical infrastructure itself.

Trust therefore became an operating property rather than something CECIDS could assume because stakeholders had participated in planning or governance. As source systems changed, staff turned over, analytic questions evolved, and products exposed new assumptions, the system had to continue earning confidence through responsiveness, transparency, and correction.

The strongest evidence of trust was not the absence of disagreement. It was the existence of a reliable way to surface concerns, examine evidence, make a decision, and follow through.

9.2 Making Reliability Visible

As CECIDS learned from these experiences, it made its reliability more visible through a set of increasingly deliberate practices. Acknowledging receipt of a file and summarizing what the system believed it contained created an early opportunity to identify a mismatch. Reconciling material differences with data-contributing organizations helped distinguish genuine errors from differences in definitions, periods, inclusion rules, or operational practice. Documenting the resolution preserved what had been learned and reduced the likelihood that the same uncertainty would have to be reconstructed later.

Reliability also required candor about what the system did and did not know. Estimates needed to be distinguished from observed counts. Assumptions and limitations needed to be stated. When timelines or conditions changed, early communication and a credible revised path mattered more than maintaining an unrealistic commitment. When an error occurred, correction and explanation protected credibility more effectively than defending an output because substantial effort had already been invested in it.

Closing the feedback loop was equally important. Stakeholders who contributed time, expertise, or data needed to be able to see what happened as a result: what was incorporated, what was not, why a decision

was made, and what would happen next. Without that follow-through, participation could easily become ceremonial rather than operational. These practices are deliberately mundane, and their importance lies precisely in their repetition: trust becomes durable when reliable behavior is routine rather than exceptional.

9.3 Different Stakeholders Need Different Assurances

The same system can be trustworthy for different reasons to different stakeholders. Data-contributing organizations and program leaders need confidence that their records will be represented accurately, that discrepancies will be investigated respectfully, and that the system will not characterize their programs without understanding the underlying context. Policymakers and advocates need information whose population, period, measure, and limitations are clear enough to support the claim or decision at hand.

Researchers need documented data, transparent methods, stable definitions, and access to appropriate program or community expertise. Families and communities need assurance that individual children are protected and that aggregated reporting will not expose, stereotype, or misrepresent small populations or communities. Those assurances cannot remain abstract promises. Governance, access controls, disclosure review, aggregation rules, documentation, and responsible interpretation make them observable.

Not every stakeholder needs to participate in every decision, and trust should not be confused with universal agreement. What matters is that roles are clear, the evidence behind decisions is visible, concerns have a legitimate path for review, and the system responds consistently.

This perspective connects trust back to every earlier section of the Field Guide. Governance makes authority visible. Architecture and documentation make technical decisions traceable. Product design connects information to legitimate use. Research governance distinguishes stewardship from methodological control. Together, these practices create the conditions under which organizations can depend on shared information.

Key Principle: Trust is an operating property of the system. It grows when governance, technical practice, stewardship, and communication make reliability and accountability visible over time.

10. Sustainability and Responsible Closeout

10.1 Sustainability Beyond Platform Permanence

CECIDS' experience ultimately broadened the meaning of sustainability. Sustainability is often discussed as whether a platform, program, or funding stream can continue indefinitely. For an integrated data system, that definition is too narrow. A technical platform may change, migrate, or end while much of the infrastructure created around it remains valuable.

CECIDS produced more than integrated records. It created governance practices, decision rights, business rules, source documentation, technical logic, analytic definitions, research assets, product guardrails, and

relationships across organizations. Those assets represent accumulated institutional knowledge. Preserving them can reduce the cost of future system-building and help subsequent efforts avoid having to rediscover decisions that have already been examined and documented.

This broader view also changes how transition is interpreted. Ending or changing a platform is not inherently evidence that the work failed to endure. The more important question is whether the value created can travel: whether definitions remain understandable, technical decisions remain traceable, stewardship obligations remain clear, and future systems can build on rather than erase prior learning.

Sustainability therefore begins during system development, not at the moment of closeout. The documentation, governance, versioning, and ownership practices described throughout this guide are also the practices that make responsible transition possible.

10.2 Preserving Knowledge and Stewardship

Transitions expose weaknesses that may remain hidden during ordinary operation. If a business rule exists only in a staff member's memory, if a transformation cannot be traced to its source interpretation, or if ownership of a research asset is unclear, the risk becomes most visible when people, contracts, or platforms change.

CECIDS' documentation and stewardship practices were designed in part to reduce that dependence on individual memory. Structural information, program-year conditions, source interpretations, technical mappings, and analytic assumptions could be preserved in forms that others could use. Version history and provenance helped explain not only what a current rule was, but how and why it had changed over time.

The same principle applied to ongoing obligations. Approved research, legal responsibilities, data-use conditions, or stewardship commitments may continue after routine platform operations have ended. Responsible closeout therefore requires clarity about which obligations persist, who holds authority, what assets remain active, and how future questions will be handled.

Preservation is not the same as keeping everything indefinitely. It is a deliberate decision about what knowledge, evidence, assets, and responsibilities must remain available because they continue to carry public, legal, analytic, or institutional value.

10.3 Closeout as a Designed Process

Closeout is itself an operating phase. Just as an integrated data system requires design decisions about governance, architecture, use, and maintenance, it also requires design decisions about how work will conclude or transition. Treating those decisions as an administrative afterthought can put at risk the same stewardship and institutional memory that took years to build.

At the Field Guide level, the relevant questions are categorical rather than procedural. Who retains authority for unresolved obligations? Which agreements or responsibilities continue? What happens to data, research assets, documentation, and technical materials? What must be communicated to

participating organizations, product users, researchers, and other stakeholders? What follow-through is required after routine operations end?

Those questions should be answered deliberately and documented clearly. Detailed transfer mechanics, final data disposition, agreement termination, and archival locations are appropriately handled in a separate closeout addendum because they are specific to the final institutional arrangements. The broader lesson for the field is that closeout should be governed with the same seriousness as startup. A responsible ending protects the organizations and people that contributed to the system, preserves knowledge that remains useful, and reduces ambiguity about what comes next.

10.4 Sustainability as Capacity That Can Travel

Some of the most durable outputs of CECIDS are portable by design. A documented funding relationship can inform another data environment. A validated understanding of a provider source can support future integration. A governance distinction between stewardship and technical administration can shape a successor system. A product-development discipline centered on information needs can be used regardless of platform.

This is why sustainability is ultimately about capacity rather than continuity of a particular technical artifact. The system's most valuable assets include the ability to ask better questions, understand source meaning, preserve historical context, govern shared use, distinguish evidence from metrics, and maintain trust through visible operating practices.

Those capacities can survive institutional and technical transition when they are documented, owned, and communicated. They are also the return on investment from much of the invisible work described earlier in this guide.

Key Principle: Sustainability is not always permanence of the same platform. It is responsible stewardship of the value, obligations, knowledge, and relationships created by the system.

11. Conclusion: What Endures

11.1 CECIDS as Information Infrastructure

CECIDS began with a simple but consequential problem: Chicago's mixed-delivery early childhood system generated substantial information, but no single program, agency, or dataset could describe the system as a whole. Building shared information capacity required more than connecting records. It required a way to govern shared use, interpret source meaning, preserve history, translate questions into evidence, develop useful products, support research, and maintain trust across organizations.

Over time, those functions became increasingly interdependent. Governance established legitimate decision rights, but technical administration made those decisions operational. Architecture integrated information, but provider and agency knowledge made the information interpretable. Documentation preserved what

the system learned. Products translated evidence into forms that different users could act on. Research extended the questions the system could ask. Trust made participation and continued use possible.

Seen together, these elements are the infrastructure CECIDS built. The platform was one expression of that infrastructure, but not its full extent. The more durable legacy is the set of capacities that allowed a fragmented system to create and use shared knowledge with greater discipline.

11.2 What the CECIDS Experience Offers the Field

The CECIDS experience suggests several propositions for other integrated data efforts. Start with the decisions and information needs the system is intended to support rather than with the data that happen to be available. Treat governance, architecture, data meaning, stewardship, products, and research as interdependent parts of the same operating model. Preserve meaningful differences and uncertainty rather than creating common measures that imply more equivalence or precision than the sources support.

Design for both current use and historical interpretation. Fund the invisible work—legal clarity, business analysis, documentation, relationship management, and governance operations—that makes technical integration usable and legitimate. Build architecture and operating practices that can absorb new knowledge without losing traceability to what came before. Treat the ability to say “the available data cannot answer this responsibly” as evidence of maturity rather than as a failure of ambition.

Most importantly, keep returning to purpose. An integrated data system does not create value because it links more records, produces more dashboards, or supports more analyses. It creates value when the information it produces helps a public system understand itself more clearly and make better-supported decisions.

11.3 What Survives the Transition

That definition of value also changes how the end of a platform should be understood. The legacy of CECIDS should not be measured only by whether the same technical environment continues to operate. Governance practices can endure. Source knowledge and business rules can endure. Historical documentation, research assets, analytic guardrails, product disciplines, and institutional relationships can endure.

Those assets matter because future systems will continue to face the same underlying challenge: children and families experience an early childhood system that crosses programs, agencies, funding streams, and institutions, while the information needed to understand that experience is distributed across them. CECIDS demonstrated that shared visibility is possible, but only when integration is paired with governance, interpretation, stewardship, and disciplined use. The work of preserving those capacities is therefore part of the work itself.

Closing Principle: Systems do not end when platforms do. What matters most is what survives the transition.

12. Companion Resource Set

12.1 Technical and Practice Memos

The Field Guide is intended to provide the integrated CECIDS story and its transferable system-level lessons. The companion technical and practice memos provide greater depth in areas where responsible implementation requires more detailed definitions, methods, guardrails, or operating guidance.

Understanding Early Childhood Program Capacity: Definitions, Use Cases, and Guardrails

Explains why capacity is multidimensional and how the appropriate measure depends on the use case, time horizon, and available source evidence.

Understanding Early Childhood Funding Relationships: Attribution, Units of Analysis, and Guardrails

Explains how funding relationships should be represented without treating child associations, provider awards, site allocations, funded seats, and dollars as interchangeable.

Understanding Race and Ethnicity Data in Early Childhood Integrated Data Systems

Provides guidance on preserving source detail, documenting crosswalks, maintaining provenance, addressing missingness, and avoiding unsupported historical inference.

Understanding Source Data Through Structured Data Dialogues: Business Rules, Documentation, and Change Control in an Early Childhood Integrated Data System

Describes a repeatable approach for learning how source data are created, validating interpretations with data-contributing organizations, confirming material findings, and maintaining understanding over time.

Trust Is Infrastructure: Building and Maintaining Stakeholder Relationships in an Early Childhood Integrated Data System

Provides an operating framework for building and maintaining stakeholder confidence through visible governance, reliable communication, validation, realistic commitments, and follow-through.

Documenting Integrated Early Childhood Data: An Introduction to the CECIDS Documentation Templates

Explains how CECIDS separated working, structural, year-specific, and technical knowledge so that source meaning and institutional memory could be maintained over time.

12.2 Templates and Operating Artifacts

CECIDS also produced reusable operating artifacts that translate several of these principles into day-to-day practice. These materials are not intended as a required model for other systems; they provide examples that can be adapted to different governance structures, source environments, staffing models, and technical architectures.

Business Analyst SOP

Defines the responsibilities and workflow for translating provider operations and source-system behavior into maintained business and technical knowledge.

Provider Fact Sheet Template

Captures relatively stable structural information about a contributing organization, source system, programs, sites, and data relationships.

Annual Data Summary Template

Preserves program-year-specific conditions, changes, limitations, and interpretation issues that should not be treated as permanent structural characteristics.

Yearly Overview Template

Documents the programs, sites, relationships, and other contextual features active during a particular year.

CECIDS Governance Agreement

Provides the formal framework that established Data Governor and Stakeholder roles, decision rights, Technical Administrator responsibilities, and governance procedures.

Additional technical, analytic, and product documentation developed through CECIDS can be read alongside these resources as examples of how governance and source understanding were translated into operational practice.

12.3 Collaborative Works

Beyond Integrated Data System Design: Ensuring Mature Systems Address Information Gaps in Policy

Examines how information needs can extend beyond available administrative data and how decision-makers, researchers, and system partners can distinguish metrics from the evidence needed to support policy and practice. Collaboration with NORC at the University of Chicago.

Population and Participation Data Priorities: What Matters Most for Chicago Early Childhood Programs

Presents insights from a collaborative effort to strengthen data use cases for population and enrollment among early childhood education providers and system leaders in Chicago. Collaboration with NORC at the University of Chicago.

From Data to Information: A Toolkit for Investigating and Defining Gaps between Data and Policy Questions

Describes a case study and provides implementation materials to support mature data and analytics systems in identifying and specifying information gaps that persist even after reports, dashboards, and similar analytics have been produced. Collaboration with NORC at the University of Chicago.

From Data to Impact: A Multi-Partner Framework to Advance Early Childhood in Illinois

Shows how city, state, research, and philanthropic partners organized information-needs work, descriptive analysis, longitudinal analysis, and analytically complex research as complementary pathways for moving from data availability to policy and practice impact. Collaboration with NORC at the University of Chicago, Center for Early Learning Funding Equity (CELFE), Illinois Early Childhood Asset Map (IECAM), Illinois Workforce and Education Research Collaborative (IWERC), and Early Childhood Research Alliance of Chicago (EC*REACH).

12.4 Future and Closeout Resources

Some areas of the CECIDS experience warrant deeper treatment than is appropriate within the Field Guide. Potential future companion resources may address research governance and product development in greater procedural detail. These resources would extend, rather than replace, the system-level principles described here.

A separate CECIDS Closeout and Final Disposition Addendum may document the final institutional and technical details of transition, including disposition decisions, transfer arrangements, continuing obligations, and archival locations. Those details are intentionally separated from the Field Guide because they reflect the specific conditions of CECIDS' closeout rather than general design principles.

The Field Guide itself is intended to stand as the complete synthesis of what CECIDS learned about building, governing, using, and stewarding integrated early childhood information infrastructure.