

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

By Emily Wiegand and Leah Gjertson

This toolkit 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. While the process of obtaining, combining, and preparing data for analysis is no small feat, extending the long-term value of these investments requires regularly revisiting both the information that is available and the policy and practice questions that persist. Specifying questions for data and research can be challenging, and it is common for them to need continued refinement after initial analyses are completed. Just as the original process of defining data needs requires bringing together experts who can speak to policy and process, data, and measurement strategies, the process of refining these needs requires ongoing dialogue from the same groups.

In 2025 the Chicago Early Childhood Integrated Data System (CECIDS) team, together with researchers from NORC at the University of Chicago (NORC), undertook an effort to review the early childhood research and data needs of CECIDS stakeholders, a group that includes program managers, city leaders, advocates, and funders. The CECIDS community had clearly articulated and prioritized use cases (see inset), and there were existing products that spoke to the use cases. However, the messaging from the CECIDS community was that their data needs were not being fully addressed by the current products. Our project was intended to foster a dialogue about how those data needs could be further specified, presented in different aggregates, or communicated using different approaches to be truly actionable. We have also published [lessons learned about process and approach](#) from this project, as well as our [findings](#).



## Focal Use Cases

This pilot effort focused on four use cases, prioritized by the CECIDS Executive Committee in Fall 2023 and reaffirmed as priorities in August 2025. While there are four use cases on this list, we chose to examine them together because they seemed clustered around the same general topic: population and early childhood enrollment of young children in Chicago.

1. **How many children ages 0–5 live in Chicago?** What are their demographics and family employment (e.g. single income household) characteristics?
2. What are the **long-term population trends** in neighborhoods? How might supply and demand change over time? Where should we be recruiting?
3. **What children are eligible for each type of early childhood program?** How many families are eligible for Head Start or Preschool for All but not for childcare due to low work hours?
4. What is the **geographical distribution of services** by program type (including family day care homes), size?

## Approach and Tools



### Tool 1: Pulse Check Questionnaire

#### 1. Revisit existing questions or use cases.

We met with the CECIDS Executive Committee in the summer of 2025 to present the project. In that meeting, we reshared CECIDS' prioritized use cases and asked committee members to confirm whether these questions were still priorities and whether they were currently addressed with existing data.

In the case of our project, the committee reaffirmed that the questions continued to be priorities and committee members said their organizations were "somewhat" able to answer the questions.



### Tool 2: Inventory Template

#### 2. Inventory existing reports, dashboards, and analyses.

We collected and reviewed existing published products that seemed to speak to the priority use cases in any way. In addition to collecting products specific to Chicago, we also collected reports from other jurisdictions that looked like they could speak to these questions if they were recreated in Chicago.

The inventory gave the project team a thorough grounding in what existing work on these questions existed, what common approaches were taken, and what potential gaps might be. It provided important context for the next round of conversations, allowing us to better probe for more detail in that stage.

We identified the materials for the inventory through web searches and through outreach to the Chicago early childhood research and data community. This outreach was our initial point of connection with other researchers and data specialists and a chance to hear their first thinking about what parts of the use case questions had and had not been addressed—and where.








### Tool 3: Questions for Information Consumers

#### 3. Meet with information consumers individually.

We conducted interviews with individuals from the CECIDS Executive Committee or their designees. These interviews typically included just 3-5 people, with two representatives from our team (one to ask questions and one to take notes) together with 1-3 representatives of a CECIDS partner organization.

Our focus in these conversations was to get as specific, clear, and actionable about the data needs and gaps identified by the respondent as possible. By conducting individual interviews rather than focus groups or meetings to gather this information, we made space to hear specifics of how the relevant information would be used in each organization's specific context. (See process brief [\[link\]](#) for more information about this approach.)

We limited initial conversations to 30 minutes. While the information we collected in the first 30 minutes with each stakeholder was not exhaustive, we got a clear sense for each organization's particular focus, priorities, and needs such that we felt cleared to move forward to later stages—especially since participants would have other opportunities to react and co-interpret findings.

 <p><b><u>Tool 4: Template for Coding Initial Interviews</u></b></p>	<p><b>4. Compile early findings.</b></p> <p>After interviews were completed, we conducted a basic coding exercise where we identified existing data sources used, targeted data gaps, “North Star” questions, and other themes in the notes from our interviews. We initially used colored highlighting by topic to code themes in the interview notes. We then compiled across interviews in the provided initial template for synthesis.</p>
 <p><b><u>Tool 5: Questions for Data and Methods Experts</u></b></p>	<p><b>5. Review early findings with data and methods experts.</b></p> <p>We met with individuals from the research and data communities to check assumptions about specific targeted data gaps and ensure there were no existing sources speaking to these gaps that had simply been missed in our inventory. We also reviewed the North Star questions with this audience to hear early feedback on how these questions might be addressed.</p>
 <p><b><u>Tool 6: Questions for Small Group Co-interpretation</u></b></p>	<p><b>6. Co-interpret identified information needs in small groups.</b></p> <p>We reengaged the organizations who participated in our initial information consumer interviews in small groups by role (specifically, we assembled one small group of individuals from a policy or advocacy perspective and one of program administrators).</p>
 <p><b><u>Tool 7: Template for Coding Small Group Co-interpretation</u></b></p>	<p><b>7. Analyze additional feedback.</b></p> <p>We completed another round of coding and analysis with results from the research and data expert review and the small group co-interpretation sessions.</p> <p>For the small group co-interpretation session, we synthesized feedback on priorities, solutions, and open questions by identified data gap across all respondents and conversations.</p> <p>For the data and methods conversations, this analysis simply involved identifying key ideas and themes. For example, one theme was ideas for how to explore which families are not engaging in existing community programs and why. (There was no template for the data methods analysis.)</p>
	<p><b>8. Circulate results for comment.</b></p> <p>We shared our revised results for feedback in multiple ways, including through presentations at large group meetings and in draft written form prior to publication. We invited comment in person or in written form. Our goal was to ensure that all final products felt representative and comprehensive to the CECIDS community.</p>

## Acknowledgements

Particular thanks are due to Serah Fatani and Nick Medrano of CECIDS for their partnership throughout this project. Shannon Gultinan, Julia Dennis, and Megan Maines of NORC supported the development and publication of this toolkit. We are also grateful to the members of the CECIDS Executive Committee and representatives from the Chicago early childhood research and data communities who engaged throughout this effort. This work was made possible in part by the generosity of the McCormick Foundation.

## Suggested Citation

Wiegand, E. and Gjertson, L. (2026). *From Data to Information: A Toolkit for Investigating and Defining Gaps between Data and Policy Questions*. Chicago Early Childhood Integrated Data System.

## Tool 1: Pulse Check Questionnaire

The following questions were used in Zoom polls of the CECIDS Executive Committee.

### Poll title: “Use Cases: Data Priorities Check”

*[While sharing the focal use cases on a slide]*

1. To what extent is use case 1 still a priority for your organization?
  - Not at all
  - A little bit
  - Somewhat
  - A great deal
2. To what extent is use case 2 still a priority for your organization?
  - Not at all
  - A little bit
  - Somewhat
  - A great deal

*[Repeat for all focal use cases]*

### Poll title: “Use Cases: Data Status Check”

*[While sharing the focal use cases on a slide]*

1. To what extent are your organization’s information needs in **use case 1** being met?
  - Not at all
  - Somewhat
  - Mostly
  - Completely
2. To what extent are your organization’s information needs in **use case 2** being met?
  - Not at all
  - Somewhat
  - Mostly
  - Completely

*[Repeat for all focal use cases]*

Follow up question (chat storm): Where and how are these information needs getting addressed?  
Where and how are they not getting met?

## Tool 2: Inventory Template

\*To receive a copy of this template in Excel, please contact Emily Wiegand ([wiegand-emily@norc.org](mailto:wiegand-emily@norc.org)).

The table below describes the different fields of this template and their intended contents.

Field Name	Description
Product	Product name or title
Authors	Product authors
Organization	Organization completing or publishing product
Publication Date	Date of original publication. May note frequency for recurring publications or releases.
Publisher Type (research, advocacy, nonprofit, government)	Classifies type of publisher for context. Note that some publishing organizations may be associated with multiple types.
Link	Hyperlink to materials with online presence
Description	A short description of the product, often copied from an executive summary, abstract, or other overview.
Intended Audiences (policymakers, researchers, advocates, nonprofits, public, corporate)	List of audiences to whom the product appears to be addressed. This helps to provide context on the language, accessibility, and focus of the product.
Time Frame of Data	Time frame for data used in the analysis.
Geography of Data	Geography of data used in the analysis.
Data Source	Description of data sources used in analysis.
Relevant Measures	List of measures included in the analysis with relevance to the identified set of use cases.
Subpopulations	List of subpopulations or subgroups for whom analyses have been completed.
Addresses First Use Case: What are the long-term population trends in neighborhoods? How might supply and demand change over time? Where should we be recruiting?  [Additional rows or columns for all other focal use cases.]	Free text description speaking to how well and in what ways the product speaks to each use case, with attention paid to identifying any potential gaps or aspects of the use case that the product may not address (e.g. certain time frames or subpopulations, particular measurement details or assumptions, etc.)
Limitations	Takeaways about limitations for this product across areas. May speak to accessibility, contents, relevance to use case, coverage of underlying data, etc.
Notes/Implications for Planning	Note other considerations that may impact how the project team thinks about the focal use cases and the relevance of this product.



## Inventory Template

	Product 1	Product 2
Product		
Authors		
Organization		
Publication Date		
Publisher Type (research, advocacy, nonprofit, government)		
Link		
Description		
Intended Audiences (policymakers, Researchers, advocates, nonprofits, public, corporate)		
Time Frame of Data		
Geography of Data		
Data Source		
Relevant Measures		
Subpopulations		
Addresses First Use Case: What are the long-term population trends in neighborhoods? How might supply and demand change over time? Where should we be recruiting? <i>[Additional rows for other focal use cases.]</i>		
Limitations		
Notes/Implications for Planning		

## Tool 3: Questions for Information Consumers

These questions and notes were used to guide our first round of in-depth questions with information consumers. These 30-minute conversations began with brief introductions, an overview of the project motivation and goals tailored to the individual audience and their prior exposure to this project, and then a review of the focal use case questions.

Conversations were far-ranging and did not stick to a strict interview protocol. However, we drew from these questions to guide the conversation and probe for detail.

### Core Questions (used in most or all conversations)

1. Tell me about how you think about these questions. Where do they come up in your work? How would you word the question? What is the day-to-day need for this information? *[Allow respondent to narrow focus on 1-2 use cases if that is the natural flow of conversation]*
2. Do you already use data to address some of the questions we've discussed? How? Where do you get the data? What do you do with it?
3. Where are there gaps between what you are able to answer with the data you have, and what you really want to know? Which gaps are the highest priority to address?

### Supplemental Questions (used occasionally as applicable and as time allowed)

4. Are there others at your organization who might have additional answers? Or are there any other ways you might think of or use these questions beyond what you've already said?
5. (If interviewer has ideas for products that may speak to an identified need close at hand) Take a look at this product with me. How well does this speak to your need? What would you need from this product (different populations, time periods, more readable, etc.) to make it something you could use within your organization?
6. What do you think is the biggest barrier to your organization having the data you need to speak to this use case? (e.g., data do not exist, lack of analytic capacity, lack of research translation/interpretation, data that exist aren't for the right subpopulations/measures, etc.)
7. What would your dream product look like and what would you do with that data?

## Tool 4: Template for Coding Initial Interviews

\*To receive a copy of the Excel version of this template, please contact Emily Wiegand ([wiegand-emily@norc.org](mailto:wiegand-emily@norc.org))

The following four tables were used to organize notes by topic across conversations after our initial round of interviews with the CECIDS community. The colors in the table names reference the highlighting we used in interview notes by topic when we initially coded those notes.

The tables below describes the set up and contents of each table.

**Table 1 Contents: Blue - Data Sources Used**

Column Name	Description
Organization	Responding organization (this spreadsheet is one row per interview)
[Columns for each data source]	One column for each identified data source, with an "X" in the relevant column for any data source cited as currently being used to speak to these questions in the organization's interview. This format allowed us to quickly look at how many organizations referenced a particular data source (columns) as well as how many data sources a particular organization referenced (rows).
Other	Text notes for other data sources referenced by the organization beyond those with specific columns. This was used for data sources that were only referenced in one conversation.
Notes	Any other detail about which data are used or how that is helpful to know

**Table 2 Contents: Yellow - Targeted Gaps**

Column Name	Description
Interviewee	Organization where this gap came up (if multiple organizations identified the same or similar gaps, those are all listed here to emphasize the extent of the demand)
Gaps	Specific data gap identified in whatever detail is available or makes sense (for example, "Data on language spoken in the home" or "Information on program catchment areas (home visiting)". Each gap gets its own row so there are multiple rows for each organization/interview. Note that this list was as inclusive as possible of potential data gaps, but some were later identified as subsets of north star questions instead of targeted data gaps for project results.
Why	Notes describing why the data described would be useful and how they would be used. Keeping this context with the ask ensures that the actual need doesn't get obscured while being summarized.
Notes	Thoughts from the internal team about the gap, how it might be met, or other dimensions to consider.

### Table 3 Contents: Pink - Big Questions

Column Name	Description
Interviewee	Organization where this question came up (if multiple organizations identified similar questions, those are all listed here separately)
Big Question	Broad questions/information needs identified by each organization (for example, "Where should we be attending events or otherwise recruiting to reach under-enrolled populations?", "Are current salaries/salary floors for ECE workers adequate?", "Would higher salaries improve retention?"). Each question gets its own row so there are multiple rows for each organization/interview.
Related North Star Question	As we developed the North Star questions during our analysis, we used this column to note the number of the North Star question that included each of these big questions. That allowed us to track if any of these big questions identified in the interviews were not adequately captured by our list of North Star questions.

### Table 4 Contents: Green - Other Themes

Column Name	Description
Interviewee	Organization where this comment or theme came up
Comment	Note or excerpt from interview of interest that was not captured in the rest of the coding scheme
Potential Theme	Coder thoughts on potential theme that may be represented in this comment. In our project, we had some additional theme content around data literacy and communication around data.

**Table 1: Blue - Data Sources Used**

	Organization	Data Source 1	Data Source 2	Data Source 3	Other	Notes
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						

**Table 2: Yellow - Targeted Gaps**

	Interviewee	Gaps	Why	Notes
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				

**Table 3: Pink - Big Questions**

	Interviewee	Big Question	Related North Star Question
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

**Table 4: Green - Other Themes**

	Interviewee	Comment	Potential Theme
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

## Tool 5: Questions for Data and Methods Experts

These questions were used to guide our initial co-interpretation of early findings with data and methods experts. These 30-minute conversations began with brief introductions, an overview of the project motivation and goals tailored to the individual audience and their prior exposure to this project, and then a review of the focal use cases. We then shared some background on the interviews with information consumers, including a summary of the participating organizations. We also shared findings about what resources were being used.

The primary purpose of these conversations was to present the targeted data needs and North Star questions initially identified and learn how experts would approach addressing those needs. As with our other conversations, we did not stick to a strict interview protocol but drew on these prepared questions.

*[while sharing targeted data needs or North Star questions]*

1. Have you done work that addresses any of these needs?
2. What other methods or existing data could meet these needs?
3. Are there opportunities to guide data users towards a different definition or conceptualization of the data need?
4. What is the ease or difficulty of implementing the methods or measurements we've discussed today? Are there any "low-hanging fruit" to consider?
5. How would you prioritize potential next steps based on likely impact?

## Tool 6: Questions for Small Group Co-interpretation

These questions were used to guide our co-interpretation of early findings with role-based small groups. These 30-minute conversations began with brief introductions, some summary information about who participated in the initial conversations, and a review of the focal use cases. We shared findings around what data sources participants were currently using as well as early versions of the North Star questions. Because these sessions were short (just 30 minutes), we focused the available time on a preliminary list of targeted data gaps. We shared a slide with seven numbered data gaps and structured the conversation off the following questions. Our moderator used prompts to encourage turn-taking and ensure all respondents had a chance to contribute on each question.

*[while sharing targeted data needs]*

1. Is there anything on this list that surprises you or that you don't understand?
2. Which of these do you see as the most important to address or that will have the biggest impact on your organization? Why?
3. Do you have any ideas for how to meet one or more of these gaps?

## Tool 7: Template for Coding Small Group Co-interpretation

*\*To receive a copy of the Excel version of this template, please contact Emily Wiegand ([wiegand-emily@norc.org](mailto:wiegand-emily@norc.org))*

The following three tables were used to organize notes by data gap across participants after our small group co-interpretation sessions with the CECIDS community. The colors in the table names reference the highlighting we used in interview notes by topic when we initially coded those notes.

The tables below describes the set up and contents of each table.

### Table 1 Contents: Yellow - Prior Gaps

Column Name	Description
Gap #	This number corresponded to the number given to each gap on the slide we shared with respondents during our small group discussion. In the conversation, we commonly referred to gaps by number (e.g., "I think #3 is my top priority").
Identified Gap	Full text description of the identified data gap (identical to the version on the slide). Note that this table is one line per data gap.
Priority Votes	Count of total people who identified this gap as a top priority. Individuals expressed a desire to prioritize anywhere from 1-4 gaps; we did not limit the number or priority votes per person.
Priority Voters	Names of the individuals who expressed this gap was a top priority. This was helpful in synthesis, as we could see if the gap was particularly valuable to individuals in certain roles.
Notes	Any other context from our conversation about where and how this gap is a priority.

## Table 2 Contents: Green - Solutions

Column Name	Description
Gap #	This number corresponded to the number given to each gap on the slide we shared with respondents during our small group discussion. In the conversation, we commonly referred to gaps by number (e.g., "I think #3 is my top priority").
Identified Gap	Full text description of the identified data gap (identical to the version on the slide). Note that this table is one line per data gap.
Solutions or Ideas to Address	Solutions or ideas about how to fill this gap sourced from these co-interpretation sessions. Ideas are summarized across respondents at the level of the identified data gap.

## Table 3 Contents: Pink - Questions

Column Name	Description
Gap #	This number corresponded to the number given to each gap on the slide we shared with respondents during our small group discussion. In the conversation, we commonly referred to gaps by number (e.g., "I think #3 is my top priority").
Identified Gap	Full text description of the identified data gap (identical to the version on the slide). Note that this table is one line per data gap.
Surprises or Additional Questions	Anything respondents noted as surprising or curious about the data gap. Ideas are summarized across respondents at the level of the identified data gap.

**Table 1: Yellow - Priority Gaps**

Gap #	Identified Gap	Priority Votes	Priority Voters	Notes
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				

**Table 2: Green - Solutions**

Gap #	Identified Gap	Solutions or Ideas to Address
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

**Table 3: Pink - Questions**

Gap #	Identified Gap	Surprises or Additional Questions
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		