

## DEVELOPING THE EARLY CHILDHOOD PARTICIPATION DATA SET

**Karla Andrew, Bryan Preston, & Kevin Dolan**

*Illinois Early Childhood Asset Map, University of Illinois Urbana-Champaign (USA)*

### Abstract

Because data are crucial to making effective decisions at every level (from the classroom to state policy), it is essential to see data as the connective pipeline running through family decisions about enrollment, programmatic decisions on service delivery, and statewide policy decisions. Our organization provides tools to present and understand early childhood services and demographics. Recently, our organization has been collaborating to design, manage, and implement a statewide longitudinal data system, which includes the Early Childhood Participation Dataset (ECPDS). The ECPDS tracks every child enrolled in several early childhood services, which allows administrators and policymakers to see how children's participation in different programs (childcare assistance to public preschool) and use of services (early intervention to special education) develop over time. It also allows analysis by geography, race/ethnicity, disability, gender, and financial status. In partnership with research groups at other universities, we are developing this system to improve decision-making in education through better data use. Our poster visualizes how the ECPDS will function while articulating specific challenges and advantages of this system.

**Keywords:** *Early childhood, longitudinal data, state systems, policy.*

---

### 1. Introduction

The current trajectory for educational data management in the United States is toward statewide data systems that track the progress of every child that participates in public education. Each state administers its own mix of programs and services, which can be state-funded, federally funded, or a combination of both. Integrated, longitudinal data systems provide a platform for evidence that inform program evaluation, equity, and fairness. Policy decisions based on more granular data can result in better outcomes for children. Over the past several years, the development of the Illinois Longitudinal Data System (ILDS) has been slow but steady. The Illinois Early Childhood Asset Map (IECAM) has become the data steward to work with multiple state agencies and partners to coordinate and provide technical implementation of the early childhood component of the ILDS.

The Early Childhood Participation Data Set (ECPDS) creates new data pathways from state agencies, primarily the Illinois State Board of Education (ISBE) and Illinois Department of Human Services (IDHS). We discuss our progress in implementing the ECPDS and discuss the challenges that we and our partners have faced in this process. In short, child-level data has been developed across multiple programs by which researchers will soon be able to track anonymized data. Challenges include designing privacy protections, incomplete data, inconsistent data, technical limitations of large datasets, and versioning.

### 2. IECAM background

IECAM was founded in 2006 to provide comprehensive early childhood data to local and state agencies, other stakeholders, as well as the public, to improve outcomes for Illinois children. IECAM receives funding from ISBE and IDHS but is not part of any state agency. IECAM was developed in part to facilitate accountability and transparency in state government during the first phase of development of state-funded preschools. IECAM has functioned primarily as a data portal that acquires, cleans, processes, and shares early childhood service and demographic data. IECAM also works to provide helpdesk-like services for those with questions related to early childhood data. Most other states do not have a dedicated entity like IECAM to act as a steward for statewide early childhood data. IECAM's position adjacent to, but not within, state agencies allow its data practices to reliably persist through different political administrations.

### 3. The ILDS

Legislation passed by the Illinois state government in 2009 enabled creation of the ILDS, a statewide system that eventually will provide data to help to track the outcomes of Illinois children as the progress from preschool through postsecondary education — and as they enter the workforce. This will help guide policymakers on where to invest time and energy to most effectively improve achievements in Illinois. The state has implemented a statewide student identifier system that can be used to link student records along their academic path. This unique identifier removes the need to use personal identifiers to help track individuals. Once fully implemented, this system will track each child that participates in the public care and education programs from as early as prenatal through their entrance into the workforce.

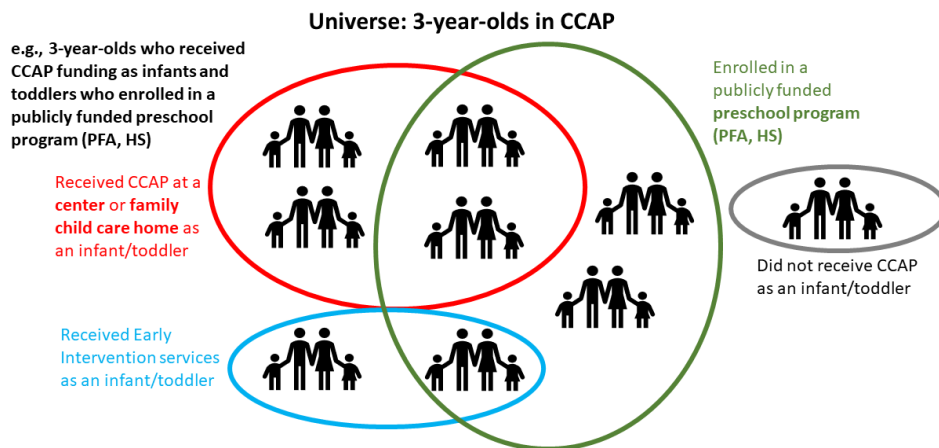
### 4. The ECPDS

The ECPDS is one element of ILDS focused on early childhood programs. The ECPDS is a massive merged dataset across agencies that can be used to track and analyze early childhood learning from birth through preschool (age 4 or 5). Because in the United States early childhood education and services are made up of a collection of disparate programs, this part of the ILDS poses organizational and logistical challenges that education data for later years (kindergarten through 12th grade) do not pose. Several services have funneled their databases into this common system, but there are more in progress. Programs already participating include:

- Preschool For All (PFA, preschool for ages 3 and 4 and some 5-year-olds; ISBE)
- Prevention Initiative (PI, mostly home-based services for children ages 2 and under; ISBE)
- Early Intervention (EI, screening and services for children ages 2 and under with disabilities or developmental delays; IDHS)
- Child Care Assistance Program (CCAP, child care subsidy for qualifying families; IDHS)
- Head Start (HS, preschool and early care services; U.S. Department of Education)

It is a uniquely scattered set of data because programs are housed under three different agencies. As a result, in the past early childhood data remained in silos with data only on the capacity to serve so many children in individual programs (PFA, PI, HS) and where (counties, cities, school districts). This system does not provide administrators and policymakers with any information on how many children have participated in more than one program or how many have been connected or “handed off” from one system to another. For example, two priority populations for PFA and HS are children living in poverty and children with disabilities, but the current system has no way to determine how many children who participated in EI or CCAP as infants and toddlers are being placed in those two preschool programs. So far, any data or information on this has been disjointed or anecdotal.

Figure 1. Tracking participation between programs.



Each of these three agencies track demographic data on children participating in their programs but do not disseminate these data widely. This has prevented advocates, policymakers, and outside administrators from evaluating whether certain demographic groups are being underserved and where. Questions often posed include:

- Are children living in poverty or children of some races/ethnicities proportionately participating in early childhood programs? Is their participation better in or worse some parts of the state?
- Are children from some immigrant communities participating in programs more than others? This could lead to the sharing of successful recruitment practices among immigrant communities.

- Are there areas/school districts where children participating in EI services are more often enrolled in publicly funded preschool programs where children can receive consistent early childhood special education services.

## 5. Implementation

Recent estimates put the population of children age 5 and under in Illinois at close to 870,000 (U.S. Census, ACS 2022). The ECPDS now contains data on all children participating in the above-mentioned programs, with monthly data for each participant extending back at least to 2015. In other words, the ECPDS is so far composed of millions of records, with millions more yet to be added.

A complex system of restricted access has been implemented to protect these data. Select Illinois state technical workers, IECAM, and their partner Chapin Hall (at University of Chicago) can access, manipulate, analyze, and repackage data in the ECPDS environment but cannot edit, download, nor share any ECPDS data. Differing levels of access are built into the design. “Curated files” will include child-level data with all pertinent demographic data available for specific agency use only. These files include relevant data such as the month and year of participation, as well as the program in which the participant took part.

For broader use by researchers, advocates, and the public, aggregated data will be accessible across the entire system and will not include any uniquely identifying data. Instead, participants are coded with a cross-program identifier that changes over time and enables tracking of individuals across programs. Aggregated data will be produced based on these identifiers. This design solves the problem of tracking individuals across programs while simultaneously ensuring anonymity.

## 6. Challenges faced in implementing the ECPDS

In addition to institutional delays, the data managers who are engaged in the process have run into several obstacles to creating the ECPDS.

- **Different reporting practices.** Some agencies report the same data differently. For example, ISBE reports Latinos only as a race while IDHS reports them as an ethnicity, preventing a direct comparison between the two agencies.
- **Geocoding children’s locations.** Critical to analyzing the participation of children in different parts of the state is being able to know where each child lives or has lived, which requires geocoding their addresses. However, the state system’s built-in geocoding system has problems geocoding addresses in some parts of the state. A more accurate system is available, but it would be cumbersome and would require downloading data off the main system, which is not allowed under project restrictions.
- **Data-sharing agreements.** An important part of the system is being able to share data between agencies and systems, which requires long, intricate agreements that often involve several rounds of sending drafts back and forth between agency administrators and, even more time consuming, lawyers. Even what seems like a minute change in an agreement can delay its approval for months.

## 7. Concluding thoughts

While this data system will soon become a vital part of the management of Illinois’s education resources, Conaway, Keesler, and Schwartz (2015) remind us that a state data system can address some of the needs for policy research, but not others. The system itself is an important component in a larger ecosystem that should also include a directed research agenda and state capacity for research. Conway et al. note that there are important questions that simply cannot be addressed by any data system, no matter how smoothly managed. For example, stakeholder perceptions about a program are not coded into state data systems. A program may appear one way with respect to what the data suggest, but that does not always indicate full workforce, or parental, endorsement. Nor can a functioning longitudinal data system provide direct answers to questions about resource allocation. For example, “should a state invest in providing direct services to its lowest performing districts, or should it instead provide those districts with grants? Should it expand after-school and out-of-school programming for the students with the greatest needs, or should it expand the school day for all students?” (Conway et al., 2015, p. 26s). Addressing such decisions may require deeper qualitative knowledge than a complex longitudinal system can provide.

## References

- Conaway, C., Keesler, V., & Schwartz, N. (2015). What research do state education agencies really need? The promise and limitations of state longitudinal data system. *Educational Evaluation and Policy Analysis*, 37(1S), 16S-28S.