

Conceptual Data Model for SLDS ECIDS

As Presented by Bardic Systems, Inc.

Table of Contents

<i>Table of Contents</i>	2
<i>Introduction</i>	3
<i>Conceptual Data Models</i>	3
<i>The CEDS Conceptual Model</i>	5
<i>The Process</i>	6
<i>Entities in the Model</i>	7
Sub-Model Persons:	7
Sub-Model Organization (and related entities):	7
Sub-Model Events and Resources:	7
Other:	8
<i>Current Draft of the SLDS ECIDS Conceptual Data Model</i>	8
Submodel: Persons	8
Submodel: Organizations	10
Submodel: Events and Resources	11
Full Conceptual Data Model	12
Data Model Requirements	13
<i>Conceptual Data Model in the Larger Process</i>	13
<i>Next Steps</i>	14
<i>Note</i>	15

Conceptual Data Model for SLDS ECIDS

Introduction

This document¹ proposes a conceptual data model for the SLDS Early Childhood Integrated Data System (ECIDS) project. It is one component in the development of a comprehensive data model for the integrated data system². As such, the model covers a wide range of data that is anticipated to be in the ideal system. The model also limits the scope of the system (by leaving things out) in order to make the first versions of the system more feasible. The conceptual model represents requirements for the eventual physical SLDS ECIDS data model.

Sub-models, organized according to a high-level Common Education Data Standards (CEDS) model, are detailed and then combined into an overall model. Also, the model is shown to align with several key indicators from the NCES ECIDS guide³.

Conceptual Data Models

A conceptual data model^{4 5}, as used in the SLDS ECIDS project, is a high-level and non-technical picture to be used by subject matter experts (SMEs) to establish the breadth of data needed for a particular system. It does not entirely

¹ Note: This paper makes use of Wikipedia articles as places to find more detail about a topic. By far, most technical Wikipedia articles are informative and accurate and serve as a good starting point for further reading. See https://en.wikipedia.org/wiki/Wikipedia:General_disclaimer.

² Wikipedia contributors. (2019, November 5). Conceptual, logical and physical schemas in: Data modeling. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:15, November 7, 2019, from https://en.wikipedia.org/w/index.php?title=Data_modeling&oldid=924709252

³ Coffey, M., Chatis, C., Sellers, J., and Taylor, R. (2014). SLDS Early Childhood Integrated Data System Guide. U.S. Department of Education. Washington, DC: National Center for Education Statistics, p. 29.

⁴ Wikipedia contributors. (2019, September 17). Conceptual schema. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:18, November 7, 2019, from https://en.wikipedia.org/w/index.php?title=Conceptual_schema&oldid=916097559

⁵ Nyerges, T. (2017). Conceptual Data Models. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2017 Edition), John P. Wilson (ed.). DOI: [10.22224/gistbok/2017.1.3](https://doi.org/10.22224/gistbok/2017.1.3)

define the system but is important in establishing the data scope of the system. This is especially important for complicated systems such as an integrated data system. The conceptual data model helps define the problems and questions that the system can address and is a way for SMEs to determine whether those goals will be met.

Conceptual data models:

- Organize the data elements into a sensible taxonomy.
- Show relationships among the items in the model.
- Make the list of data elements easy to access and understand because the list is parsed into smaller, understandable entity chunks.
- Are easy for domain experts to view, criticize, provide input, and make contributions to without needing to be conversant in database design, terminology, or tools.
- Provide a conceptual base for the database design.
- Represent data model requirements.
- Are a precursor to data normalization.
- Are a precursor to Entity-Relationship (E-R) design decisions.

The CEDS Conceptual Model

The Common Education Data Standards (CEDS) is a good resource for standardized definitions of data elements in education. CEDS is referenced throughout the data design and analysis of this project.

CEDS has promulgated a conceptual model (still in draft) that demonstrates its broad applicability across multiple domains within education. We will apply this high-level model to our SCCOE ECIDS model using the six key concepts illustrated in the picture below.



The six key concepts are:

1. Time
2. People
3. Organizations
4. Roles
5. Events
6. Resources

These key Concepts will be used to organize, at a high level, the entities in the SCCOE conceptual data model in order to make the conceptual model easier to understand. Then, the conceptual submodels will be combined into a complete conceptual model.

The Process

The identification of data elements to be included in the system and the development of the conceptual data model are important artifacts for later design of the system. An iterative process works well for development of the conceptual data model. During conceptual model development data elements are identified.

In general, the process includes the following steps:

1. An initial conceptual model is defined.
2. SMEs review and discuss the initial model.
3. The model is revised.
4. Data elements associated with entities are identified.
5. If important data elements are identified that don't fit with an entity, a new entity is created.
6. If entities have no data elements, or have a minimal number, they are dropped from the model. If an entity is deemed not important to the project, it is dropped from the model.
7. This process is repeated as many times as is practical and useful.
8. The conceptual model, plus the data elements, are used to create logical and physical models⁶.

⁶ Nyerges, T. (2017). Conceptual Data Models. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2017 Edition), John P. Wilson (ed.). DOI: [10.22224/gistbok/2017.1.3](https://doi.org/10.22224/gistbok/2017.1.3)

Entities in the Model

The entities in the model are sorted into sub-model categories consistent with the key concepts listed above⁷.

Sub-Model Persons:

- **Adult** - A fully developed person from maturity onward.
- **Child** - A young person of either sex.
- **Client** - A child or adult that receives services.
- **Parent/Guardian** - A father or mother; one who gives birth to or nurtures and raises a child; a relative, or other adult, who plays the role of guardian.
- **Staff** - The body of teachers and administrators at a school.

Sub-Model Organization (and related entities):

- **Class** - A body of students who are taught together.
- **Group** - Any number of entities (members) considered as a unit.
- **Organization** - A group of people who work together.
- **School** - A type of organization in which clients are provided education services.
- **Facility** - A location or building in which services are rendered or administration of the program is done.
- **Program** - A system of projects or services intended to meet an educational need.

Sub-Model Events and Resources:

- **Interaction** – An in-person or virtual encounter between a staff service provider and a client for the purpose of instruction or the delivery of services.

⁷ Many of the entity definitions use WordNet as their base: Princeton University "About WordNet." [WordNet](#). Princeton University. 2010.

- **Service** - Work done by one person or group that benefits another.

Other:

- **Family** – A social unit living together: includes Adults, Children, Guardians and other relations.

The Family entity is still not well defined operationally because of the many ways it is implemented in the field.

Current Draft of the SLDS ECIDS Conceptual Data Model

The model consists of:

1. **Entities** (usually represented by boxes).
2. **Lines that connect entities.** The lines are labeled with a verb clause such as: is a type of, receives, provides, etc.

An entity is a “thing” that we want to collect information about. The entity must be important enough to have data elements (or data fields) attached to it. For example, an entity called “Child” would be an important entity to include in an ECIDS, whereas Attendance is a thing but might be better represented as a series of data elements rather than an entity.

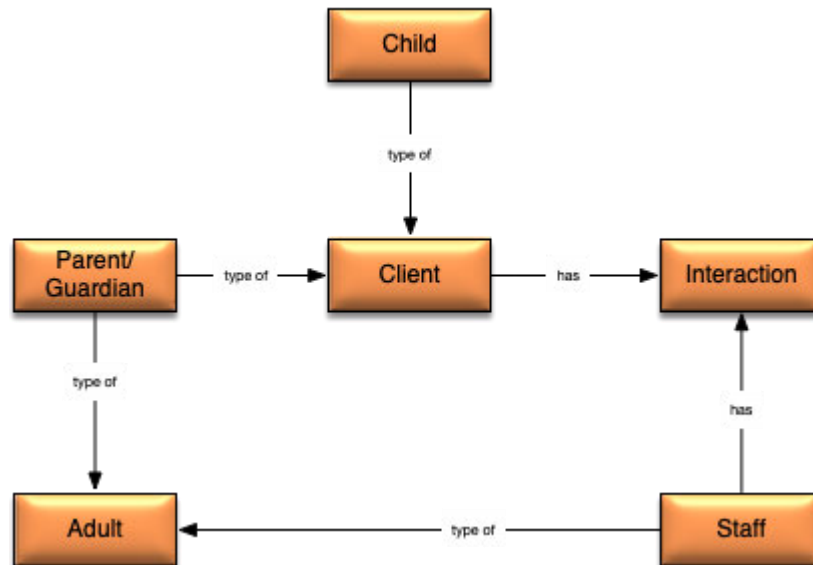
It is important to note that the arrows/lines in the diagrams below do not depict the flow of data; nor do they depict a process flow. The lines depict an operational relationship between two entities. There may be other salient relationships between two entities in the model but only those that are operationally important will be depicted in the diagrams.

Submodel: Persons

This diagram depicts mostly persons. For example, a child is a person but a child is also a type of client. The staff entity in this diagram may be considered a role rather than a person because the adult representing a staff person is in that role by association with a program or organization. The interaction entity is not a person but is shown because it is a crucial mediator between client and staff.

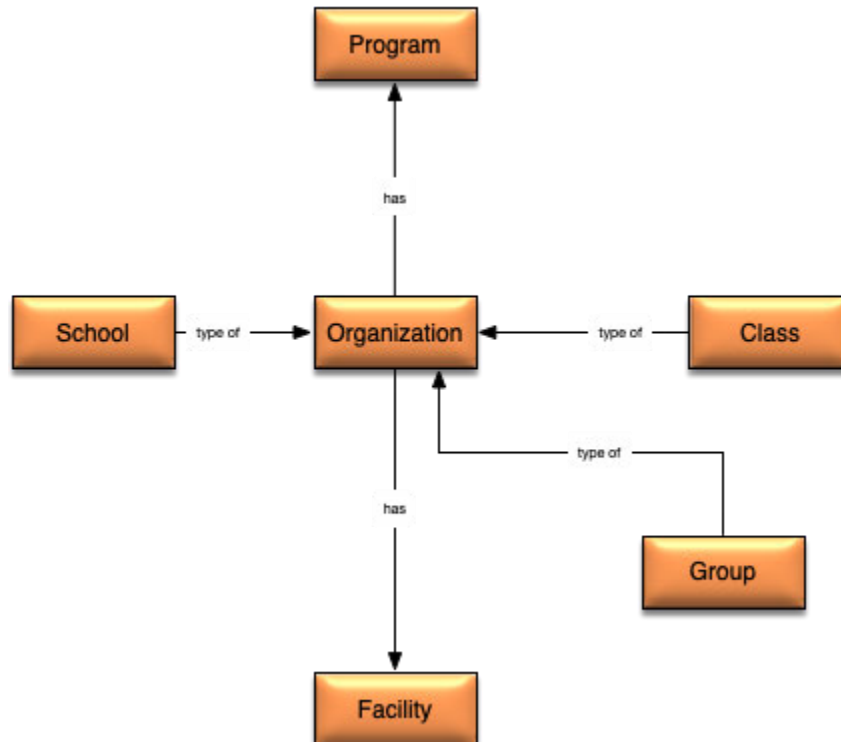
Not every possible relationship is represented in the diagram. Only relationships that are important to the requirements of the project are depicted.

For example, a staff member is usually an adult but that relationship is not shown in this diagram because it is not an important distinction for the purposes of the software system to be designed.



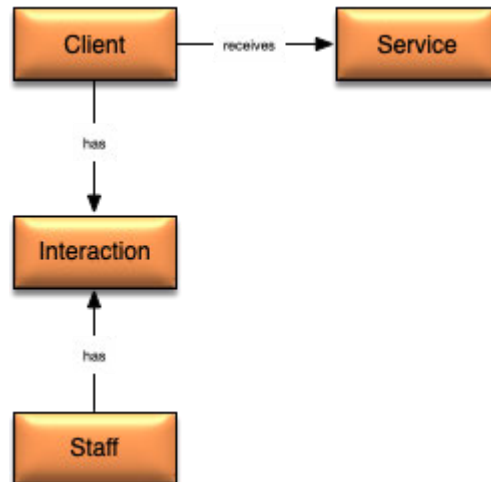
Submodel: Organizations

Organizations in this context refers to early childhood programs within the county. The diagram below shows types of organizations. However, the facility entity is not a type organization but is an important entity related to education organizations. Similarly, a program is an entity related specifically to early childhood organizations.



Submodel: Events and Resources

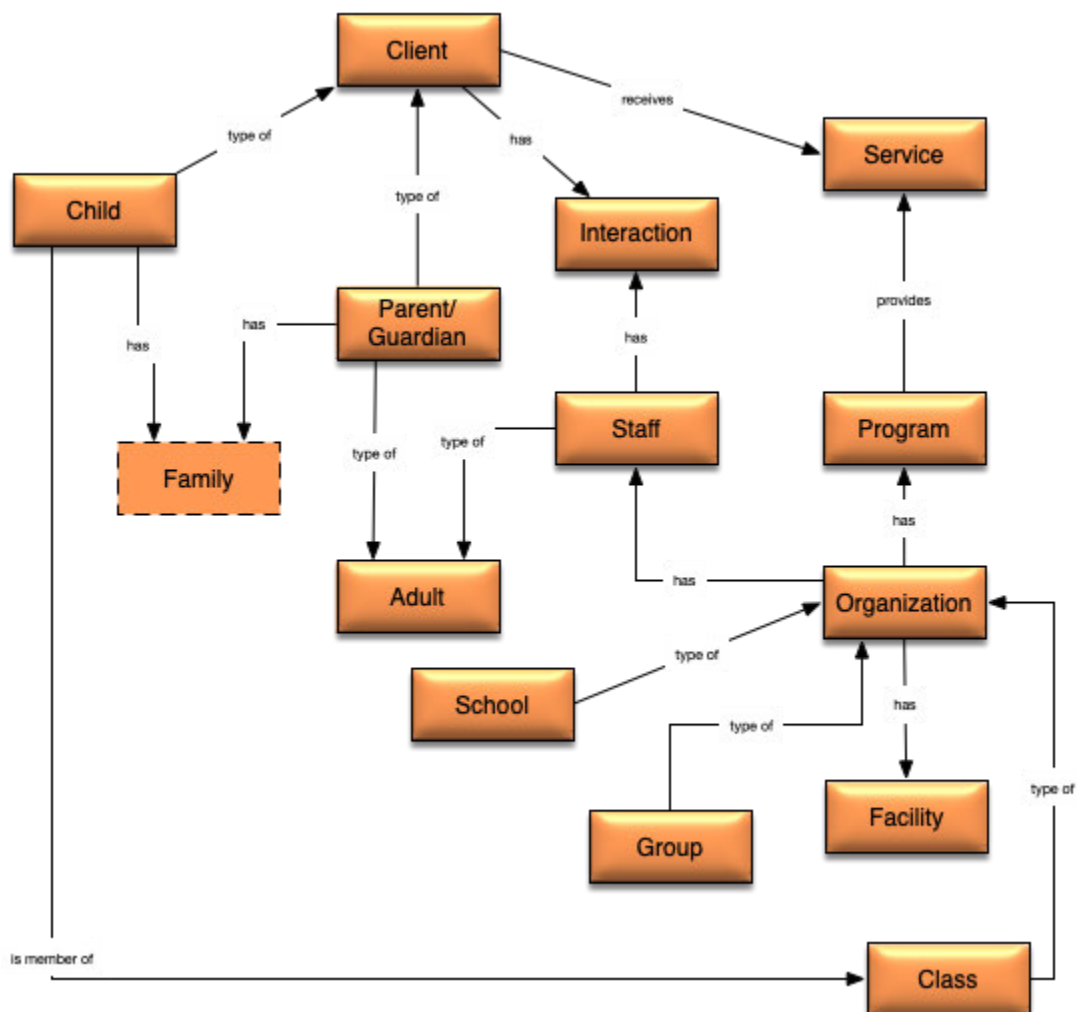
This diagram depicts the crucial point of contact between Early Childhood providers and their clients as well as the services being provided. The interaction entity lies between the client and staff. A service entity is identified as important to track.



Full Conceptual Data Model

The diagram below depicts the combined version of the conceptual data model. Although it looks a bit complicated, it is still accessible to persons who are not database technicians. One of the most important things this diagram does is establish the data scope of the project. If an entity is not in the model, then information about that thing will not be tracked, stored, or exchanged. If an entity (or relationship) is found to be irrelevant, not available, or not of interest, it should be removed now so that this type of information will not appear in the succeeding logical and physical models.

The Family entity is shown flattened and with a dotted line because how the entity should relate to other entities in the model is still in question. At this point, we believe it represents important information that should be tracked; but exactly how to implement the entity will require more research and discussion.



Data Model Requirements

As mentioned above, the SLDS ECIDS Conceptual Data Model (along with the data element list) represents requirements for the eventual data model. Here are some examples of ways that the above diagram can be translated into requirement statements:

- Both Parents and Children receive services. This suggests an opportunity for normalization during the physical data model design phase by the use of the Client entity, e.g., identifiers, names, and addresses could be kept in a single table with child-specific and parent-specific information kept separately. However, this is a design decision left for later and not mandated by the conceptual model.
- Programs provide Services. This implies that we want to track data elements related specifically to programs and services, and that there will be a relation between them.
- Staff is a defined entity in the model above and the entity has a relation that indicates interaction with children and parents. This model provides for the tracking of those interactions by specifying the Interaction entity.
- The model specifies that Services received by children and parents will be tracked with data elements.
- Organization and Facility are separate entities because an organization may have multiple facilities that need to be tracked. If this is not important then this distinction could be dropped from the model.

Conceptual Data Model in the Larger Process

The conceptual data model is a small but important part of the overall design effort for an ECIDS. The National Center for Education Statistics provides guidance for State-Wide ECIDS as part of its State Longitudinal Data Systems Grant Program (<https://nces.grads360.org/-program/ecids-toolkit>).

In part, the Conceptual Data Model addresses three key indicators in *Component E: System Design* of the document *SLDS Early Childhood Integrated Data System Guide*, page 29 (<https://slds.grads360.org/-communities/pdc/documents/8968>) provided by NCES.

The key indicators are:

- **Key Indicator 1:** The established ECIDS design meets the requirements aligned with the state's long-term purpose and vision
- **Key Indicator 2:** The system design reflects the current and continued needs for the ECIDS
- **Key Indicator 3:** The ECIDS design is articulated in a way that stakeholders, researchers, or any other nontechnical (i.e., program) team members can clearly understand the system design and its implications⁸

Although not reflected in the conceptual model, the logical model will reflect the importance of key indicator 4:

- **Key Indicator 4:** There is a unique identifier(s) (UID) or established matching process to ensure an accurate, unduplicated count of children, staff, and programs across the state⁹

In addition, a local-level Integrated Data Self-Assessment provided by NCES is being undertaken in Santa Clara County. The result of this assessment will inform the conceptual and logical data models.

Next Steps

Now that this draft conceptual data model has established a foundation and scope for the SLDS ECIDS data model, the data dictionary can be completed and the logical model constructed. The logical model will be more technical in nature and will include data elements attached to each entity. As more information is collected from Santa Clara County and other counties, this conceptual model may evolve.

⁸ Coffey, M., Chatis, C., Sellers, J., and Taylor, R. (2014). SLDS Early Childhood Integrated Data System Guide. U.S. Department of Education. Washington, DC: National Center for Education Statistics, p. 29.

⁹ *ibid.*

Note

The project described was supported by the Preschool Development Grant Birth Through Five Initiative (PDG), Grant Number 90TP0015-01-01 from the Office of Child Care, Administration for Children and Families, U.S. Department of Health and Human Services. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Child Care, the Administration for Children and Families, or the U.S. Department of Health and Human Services.